華南師範大學
SOUTH CHINA NORMAL UNIVERSITY

973项目”基于LAMOST大科学装置的银河系研究及多波段天体证认” 2016年度学术研讨会

Estimating Stellar Parameters from a Spectrum

Xiangru Li

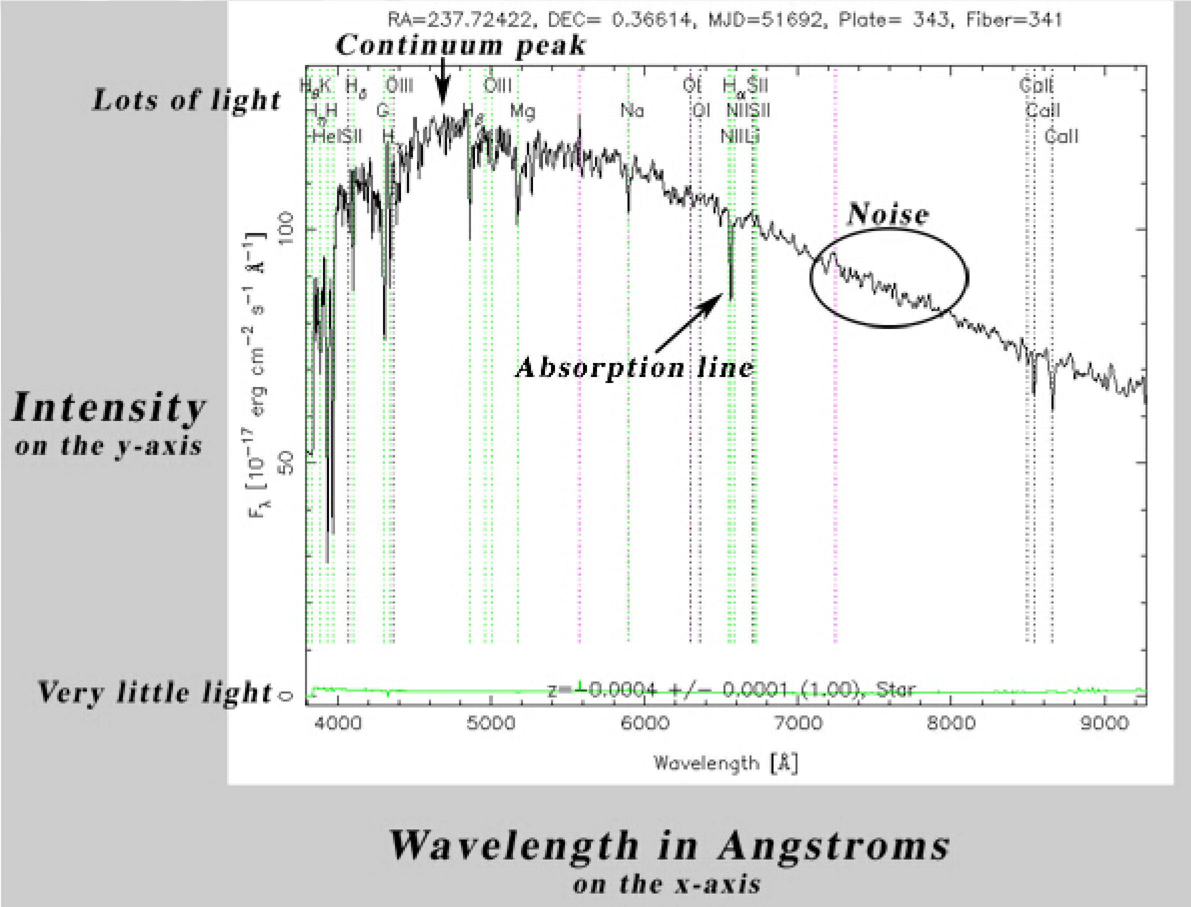
2017. 02.18

Collaborators： Ali Luo， Yongheng Zhao, Georges Comte，Fang Zuo, Q.M. Jonathan Wu, Tan Yang, Yongjun Wang, Yu Lu

Contents

- Problem, Available Schemes and Objective
- Sparse Feature Extraction
- Linearly Supporting Features Extraction

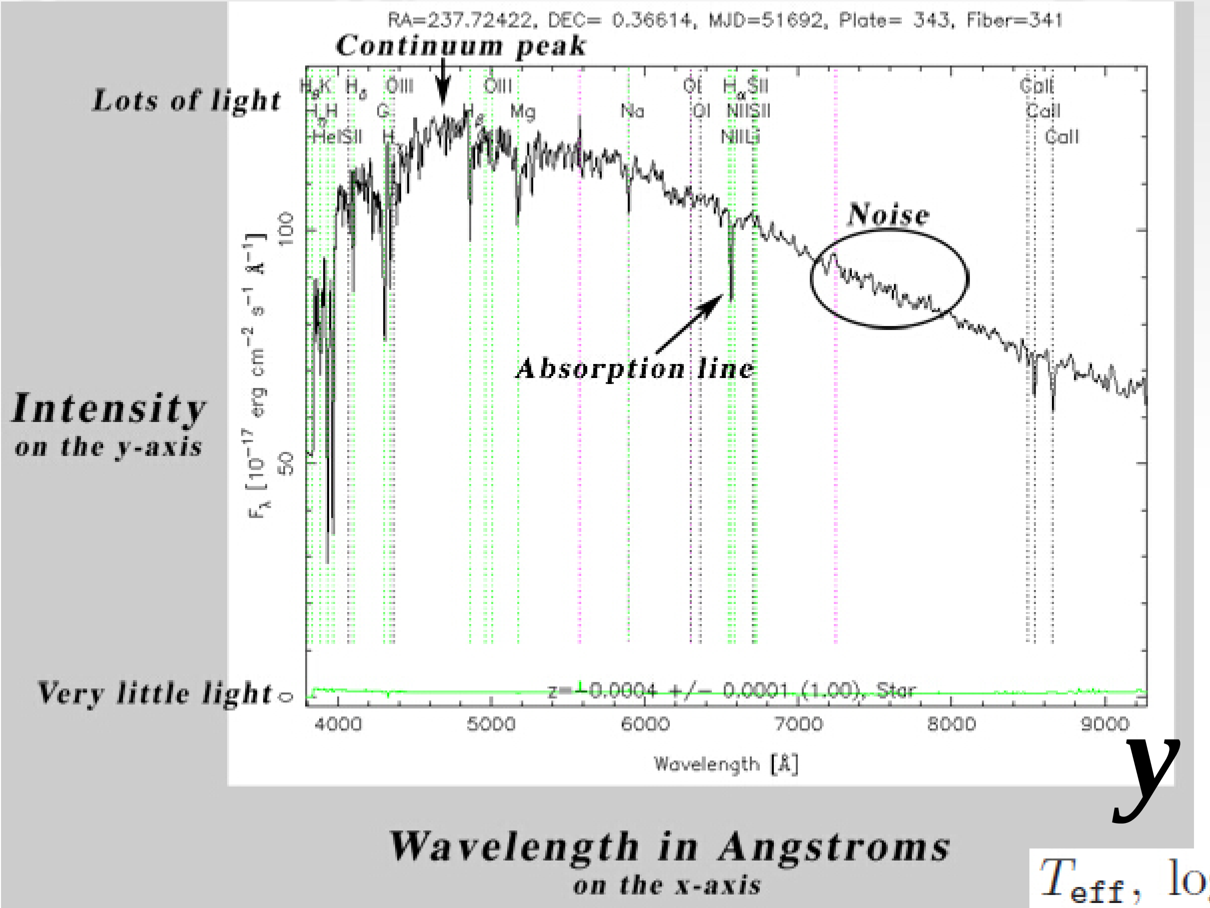
Problem



x $T_{\rm eff}$, $\log\ g$ and [Fe/H]

$$y = f(x)$$

Available Schemes and Objective



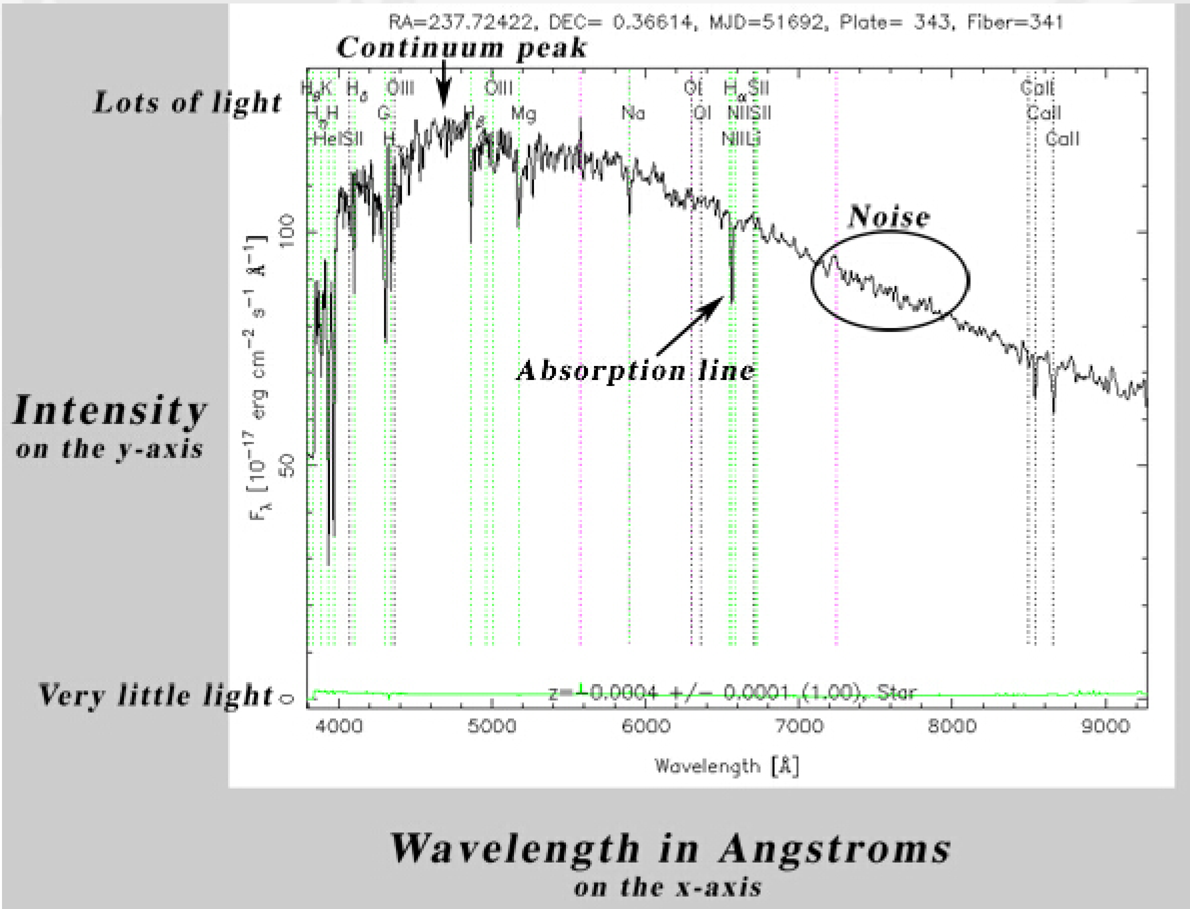
$$\boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x})$$

$T_{\mathtt{eff}}$, $\log g$ and $[Fe/H]$

- Template Matching Method
- Statistical Index Scheme(SIS)
- Line Index Method

- Physical Interpretability
- Robustness
- **local, sparse**

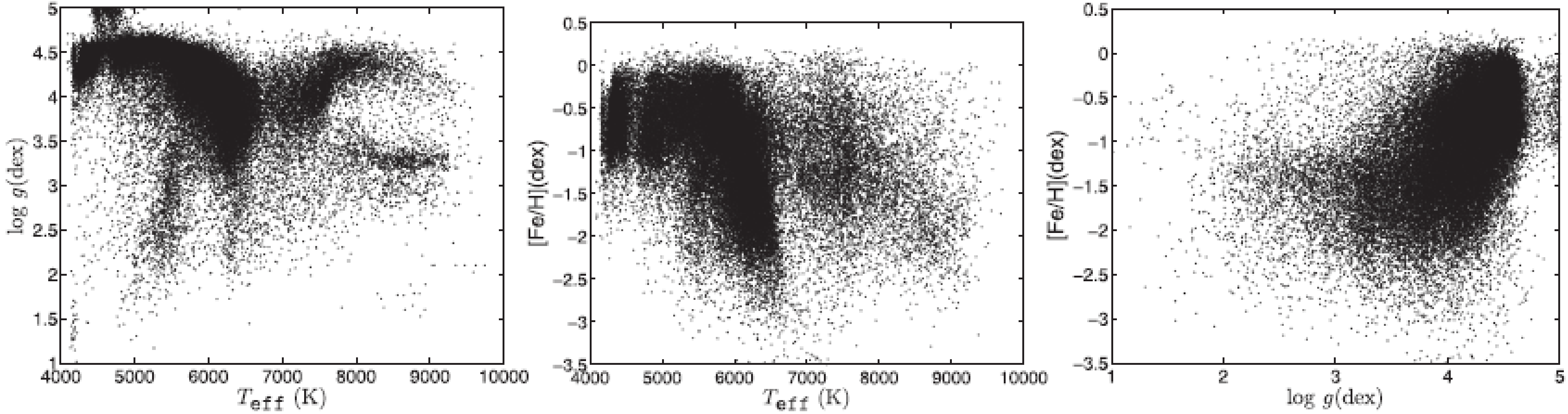
Problem and Objective



- Detection
- Description
- Estimation ← f — x

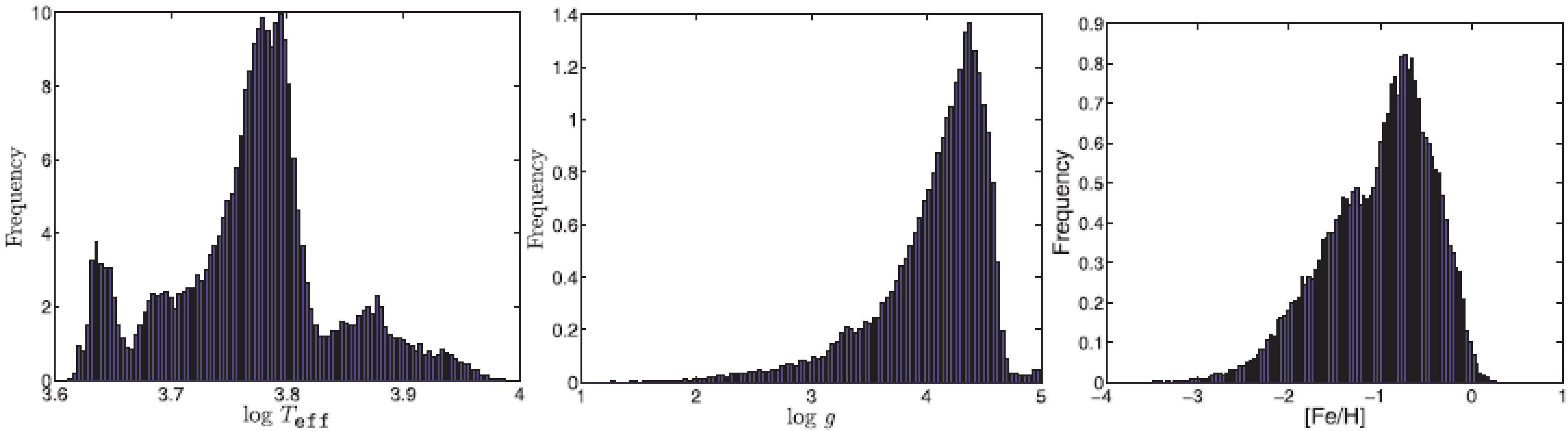
$$y = f(x)$$

SDSS Data



(a) T_{eff} and log g (b) T_{eff} and [Fe/H] (c) log g and [Fe/H]

Figure 1. Scatter diagram of the atmospheric parameters of the selected spectra.



(a) log T_{eff} (b) log g (c) [Fe/H]

50000, [4088, 9740]K for Teff, [1.015, 4.998] dex for log g, [-3.497 0.268]dex for [Fe/H]

Synthetic Data

Kurucz's NEWODF models, SPECTRUM package

18969

[4000, 9750] K for Teff, 45 values, step sizes of 100 K between 4000 and 7500, 205 K between 7750 and 9750 K

[1, 5] dex for log g, 17 values, step size of 0.25 dex

[-3.6 0.3]dex for [Fe/H], 27 values , step size of 0.2 dex between -3.6 and -1 dex, and 0.1 dex between -1 and 0.3 dex

Sparse Feature Extraction

Detection

- LASSO (least absolute shrinkage and selection operator)

$$(\hat{\alpha}, \hat{\beta}) = arg \min_{(\alpha,\beta)} \{\Sigma_{i=1}^{N} (y_i - \alpha - \Sigma_{j=1}^{p} \beta_j \tilde{x}_j^i)^2\}$$

subject to

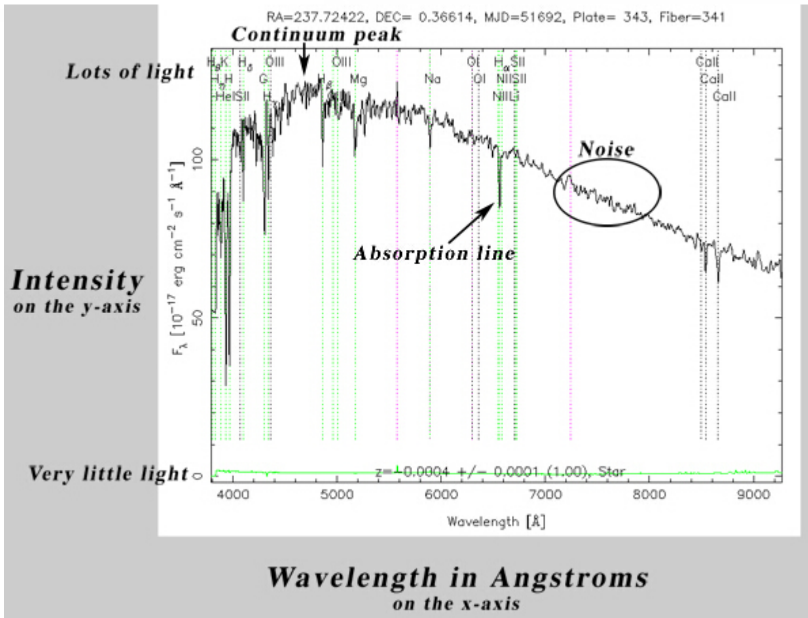
$$\Sigma_{j=1}^{p} |\beta_j| \leq t$$

where

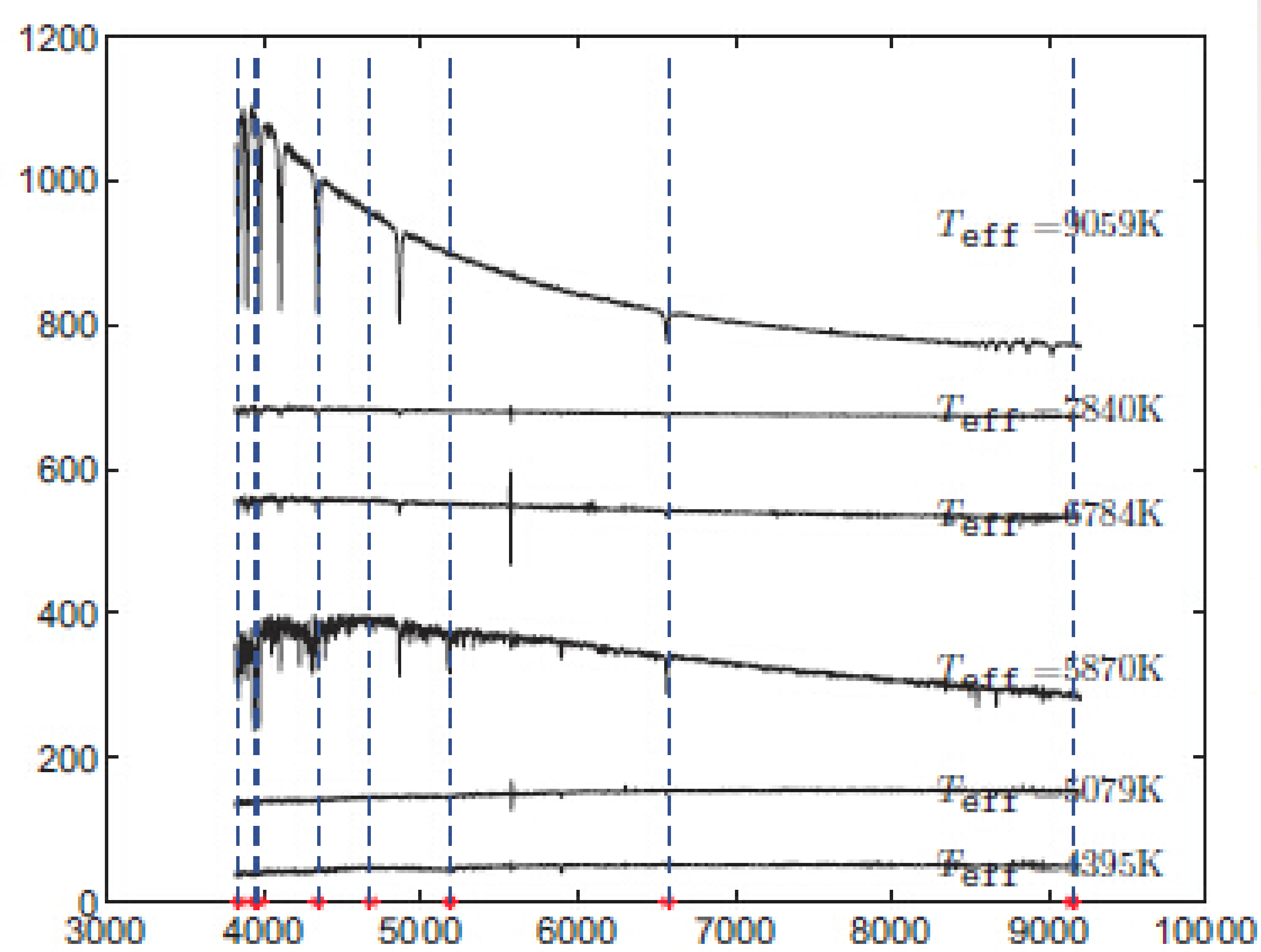
$$y_i \approx \alpha + \Sigma_{j=1}^{p} \beta_j \tilde{x}_j^i$$

$$\tilde{x}^i = (\tilde{x}_1^i, \cdots, \tilde{x}_p^i)^T$$

$$S_{tr} = \{(\tilde{x}^i, y_i), i = 1, 2, \cdots, N\}$$

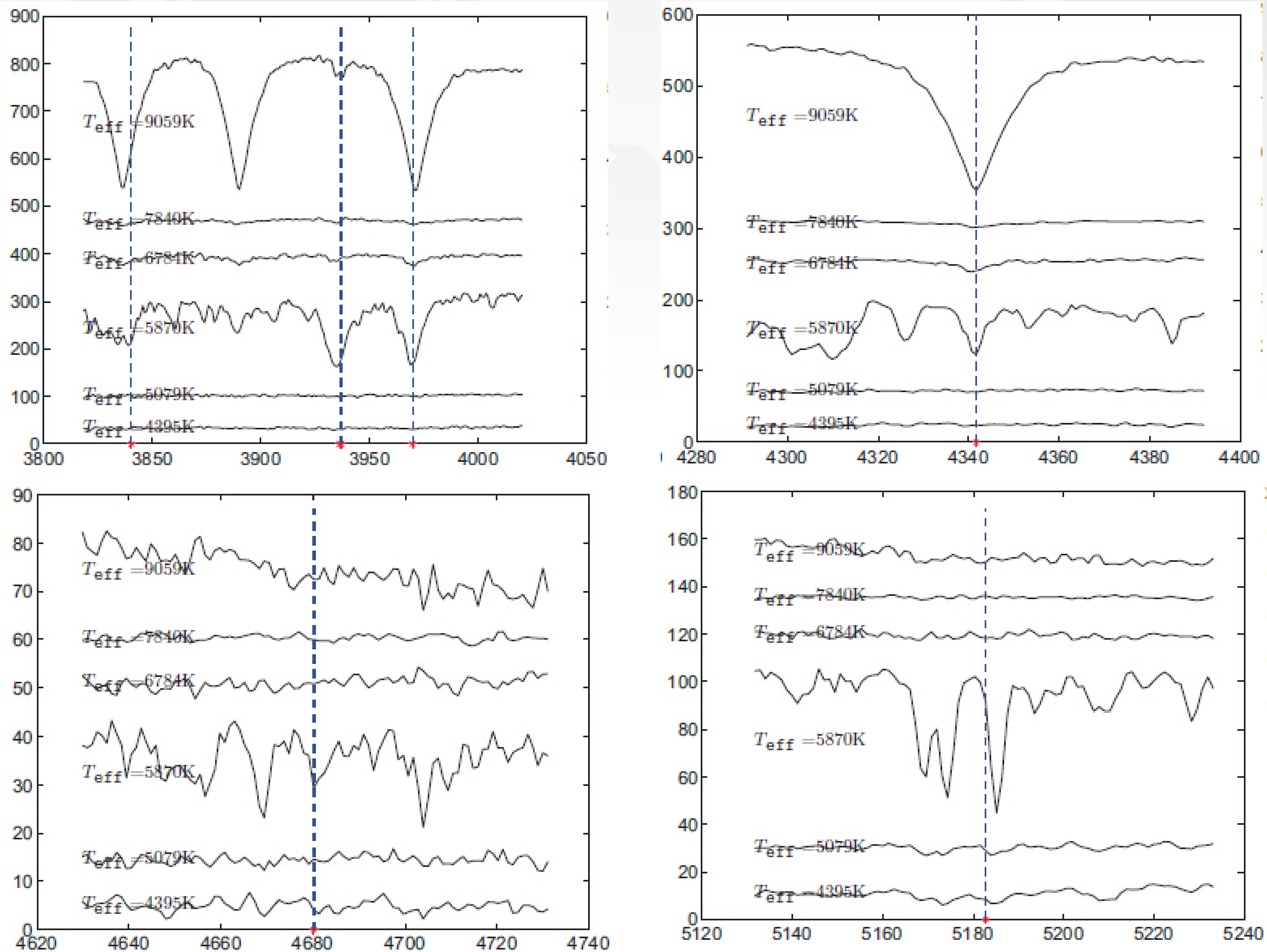


Detection



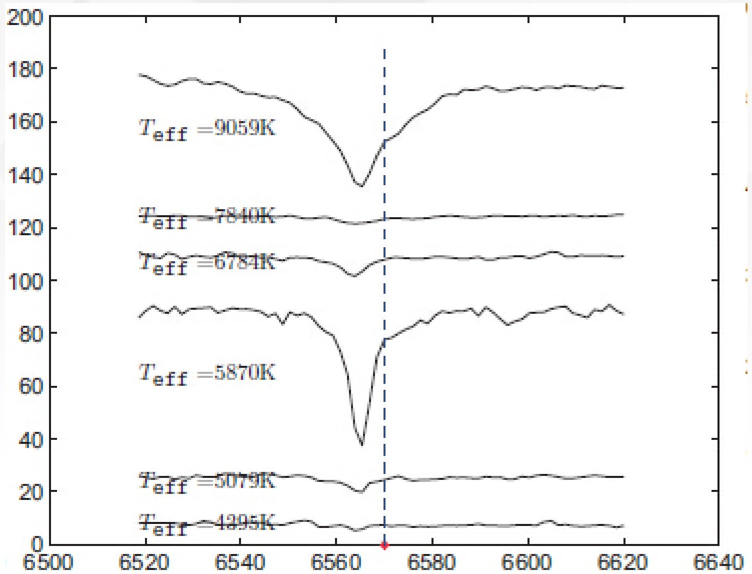
1200
1000
800
600
400
200
0
3000
4000
5000
6000
7000
8000
9000
10000
T_{eff} =9059K
T_{eff} =7840K
T_{eff} =6784K
T_{eff} =5870K
T_{eff} =5079K
T_{eff} =4395K

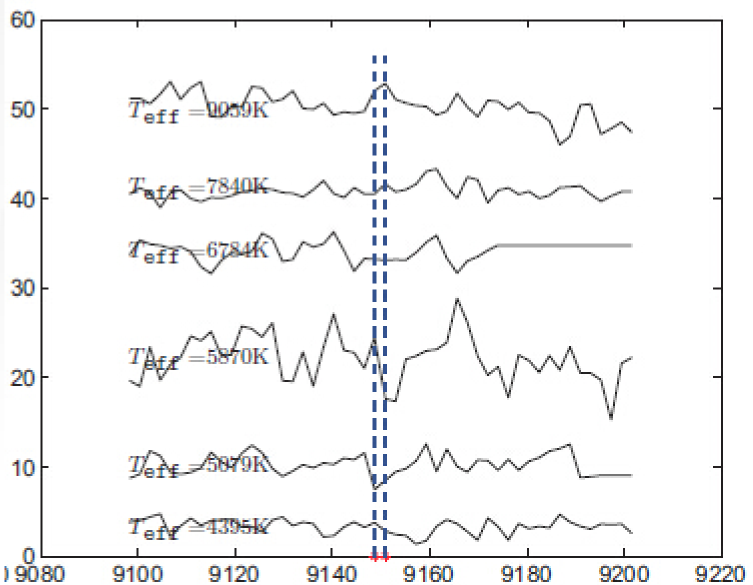
Detection



Teff =9059K
Teff =7840K
Teff =6784K
Teff =5870K
Teff =5079K
Teff =4395K

Detection





99.74%

Table 5: Consistency/Accuracy on test set with features described by the observed fluxes on the detected typical positions.

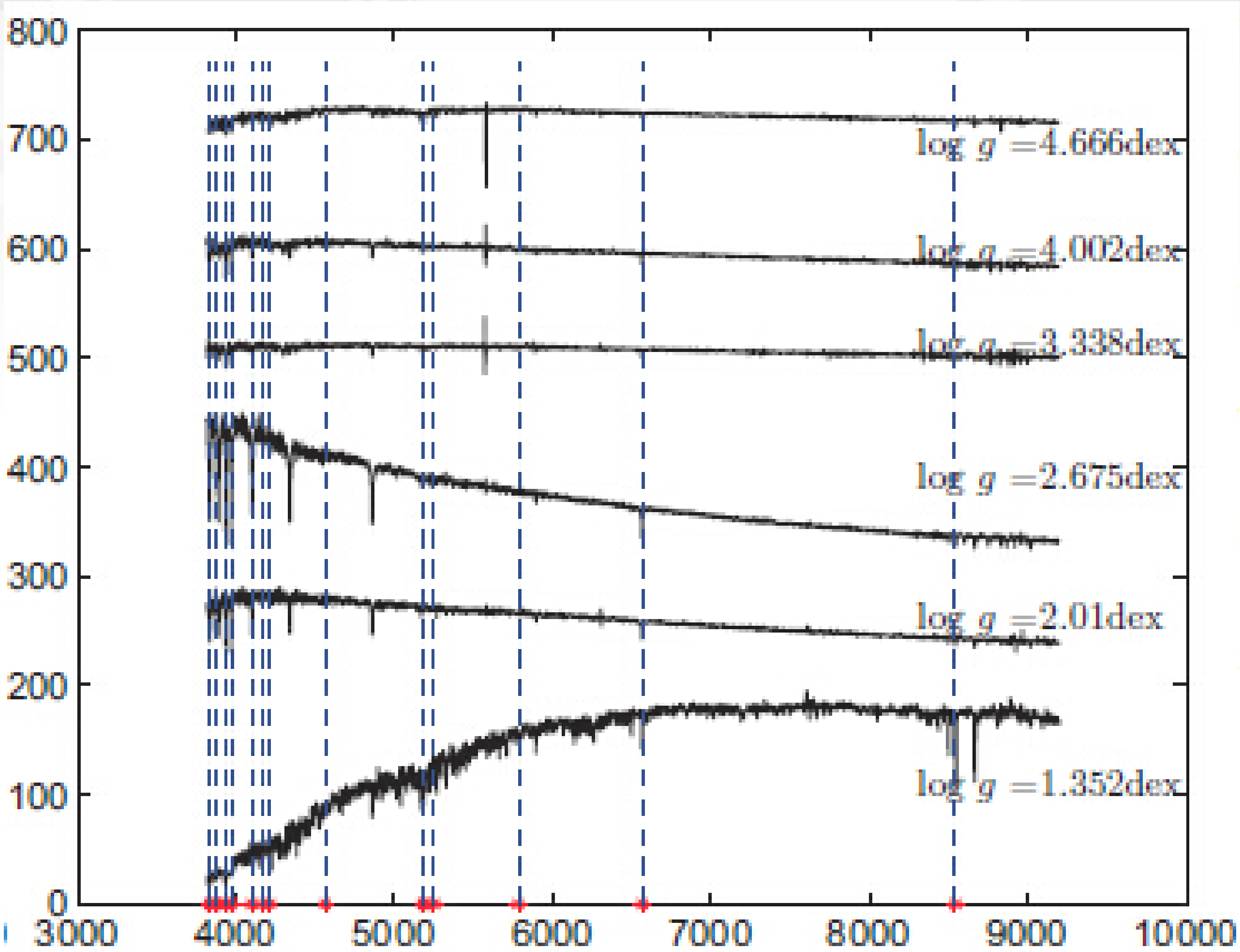
evaluation method	$\log T_{eff}$
MAE	0.009092
SD	0.012978

Table 2. Mean absolute errors on the evaluation set of 19 000 spectra in the RR model (plotted in Fig. 5).

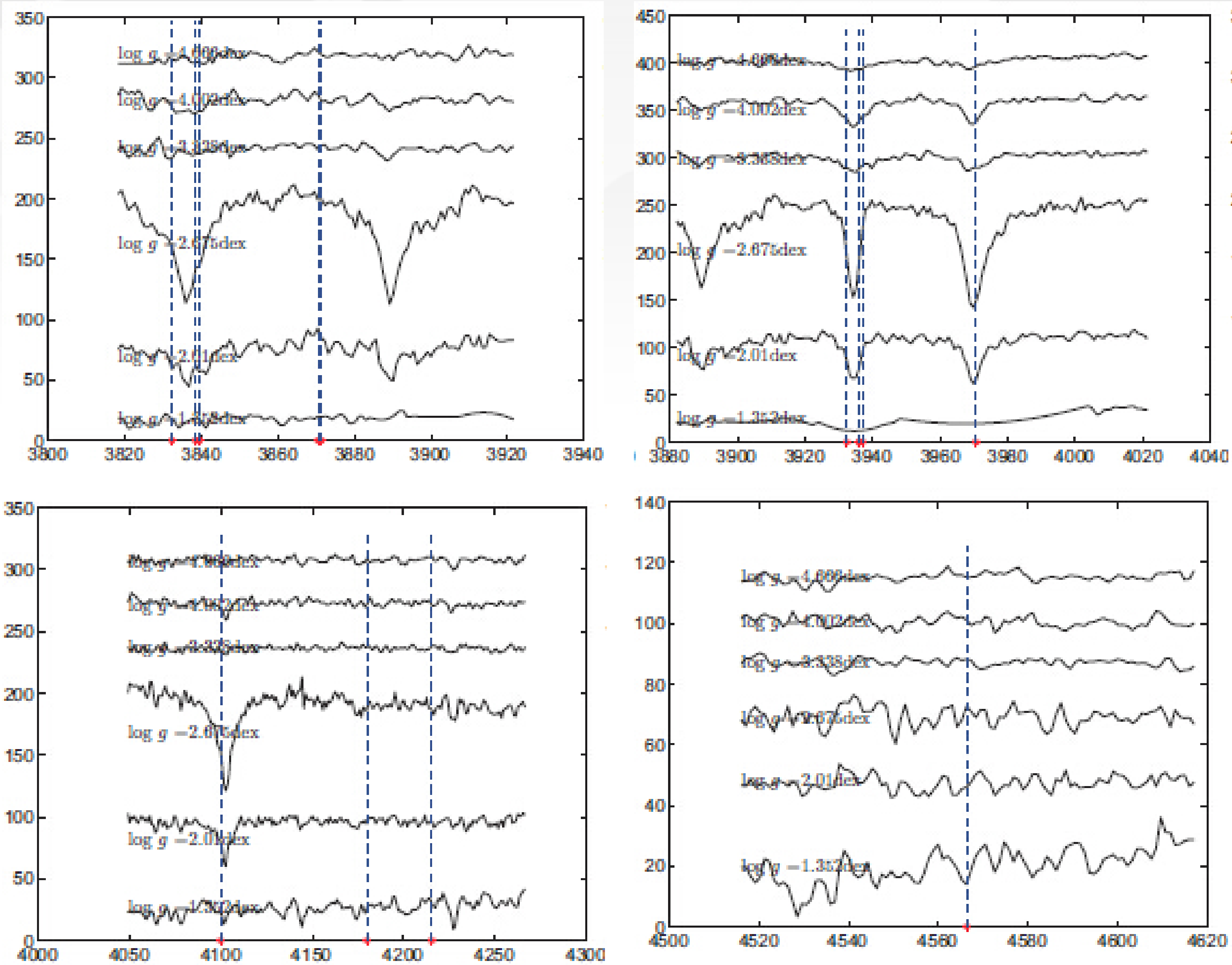
Set	PCs	+	$E_{\log T_{eff}}$	$E_{\log g}$	$E_{[Fe/H]}$
19 000	25 + 25		0.0126	0.3644	0.1949

Re Fiorentin, P., et al. 2007, A&A, 467, 1373

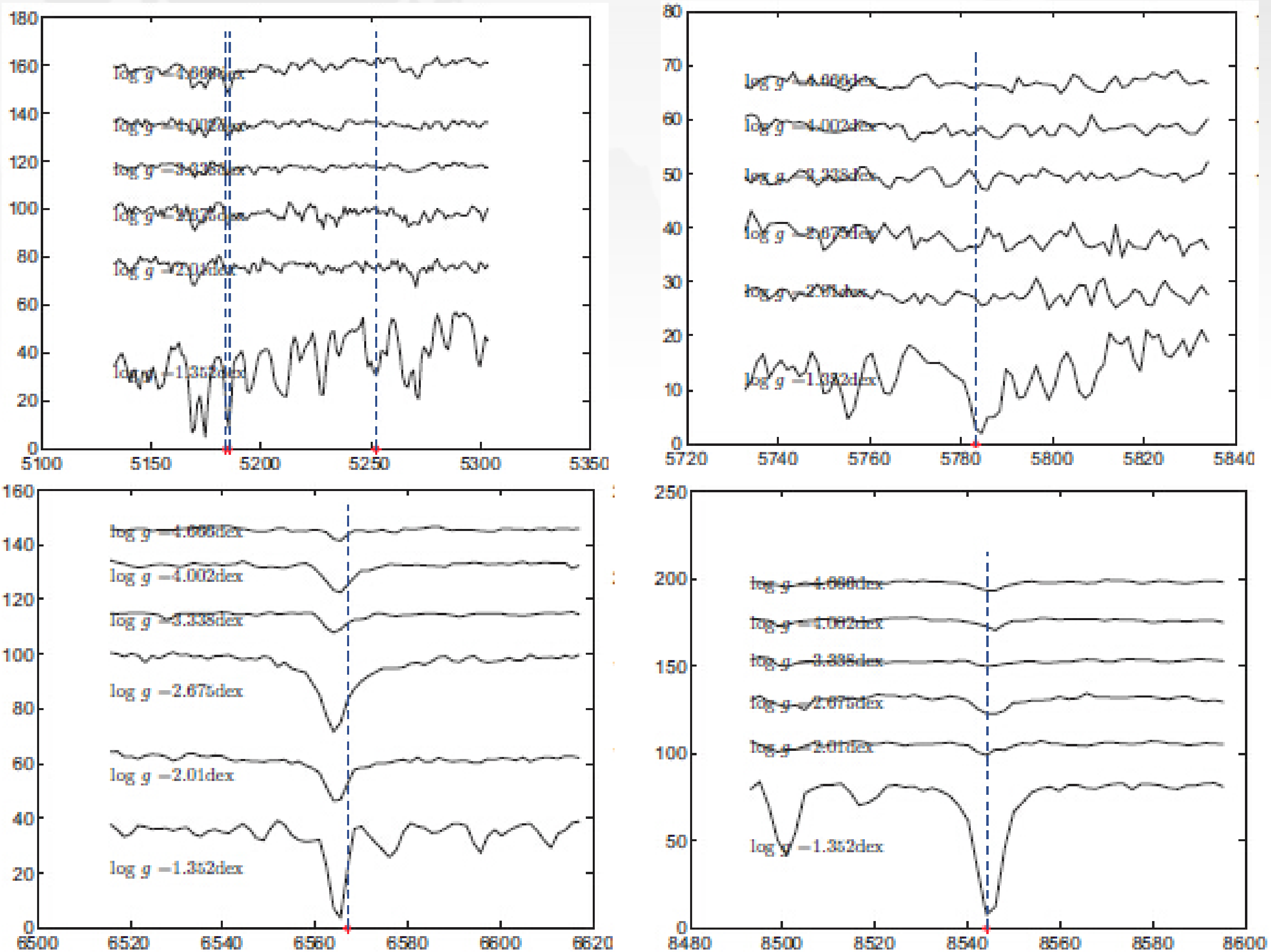
Detection


800
700
600
500
400
300
200
100
0
3000
4000
5000
6000
7000
8000
9000
10000
log g =4.666dex
log g =4.002dex
log g =3.338dex
log g =2.675dex
log g =2.01dex
log g =1.352dex

Detection

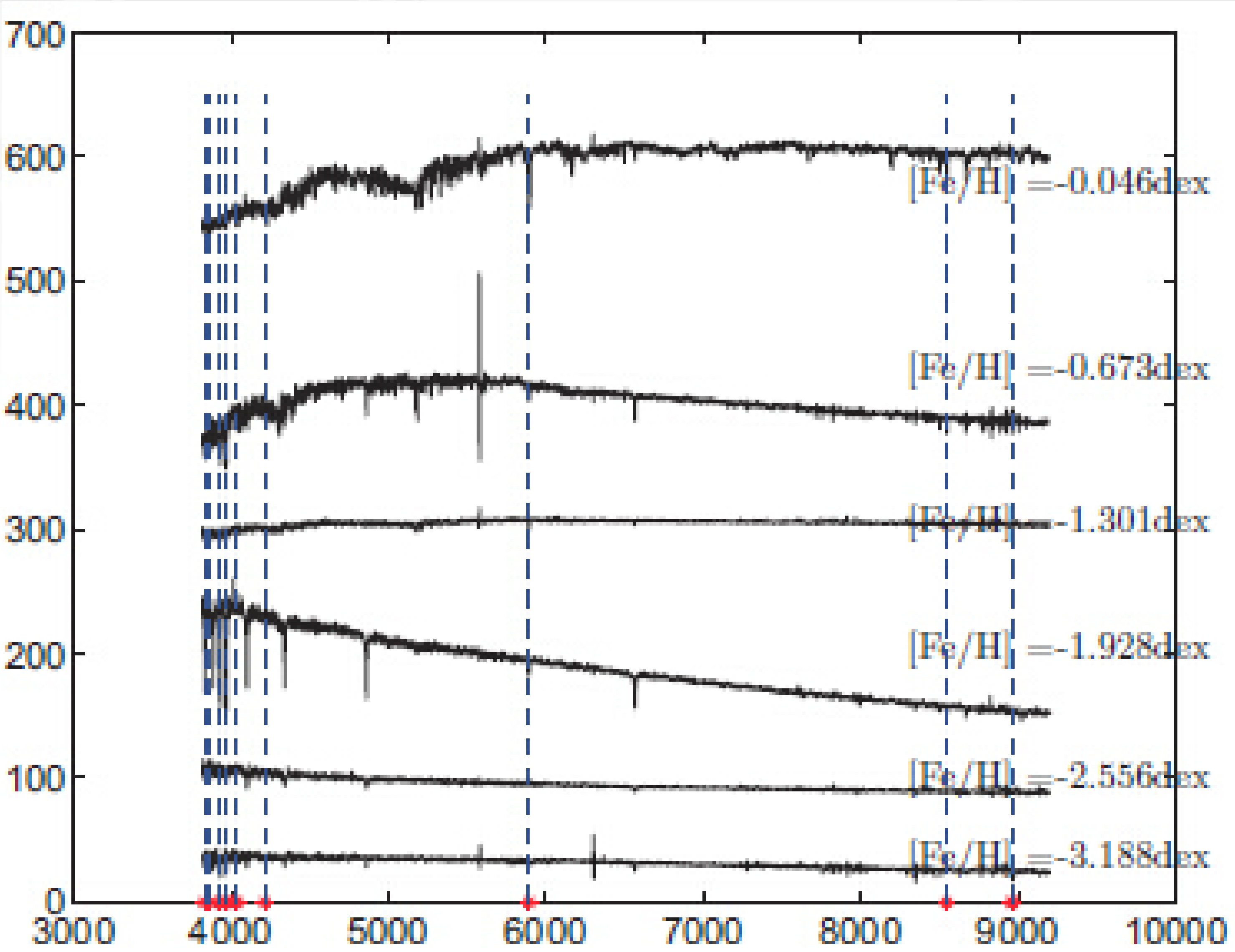


Detection



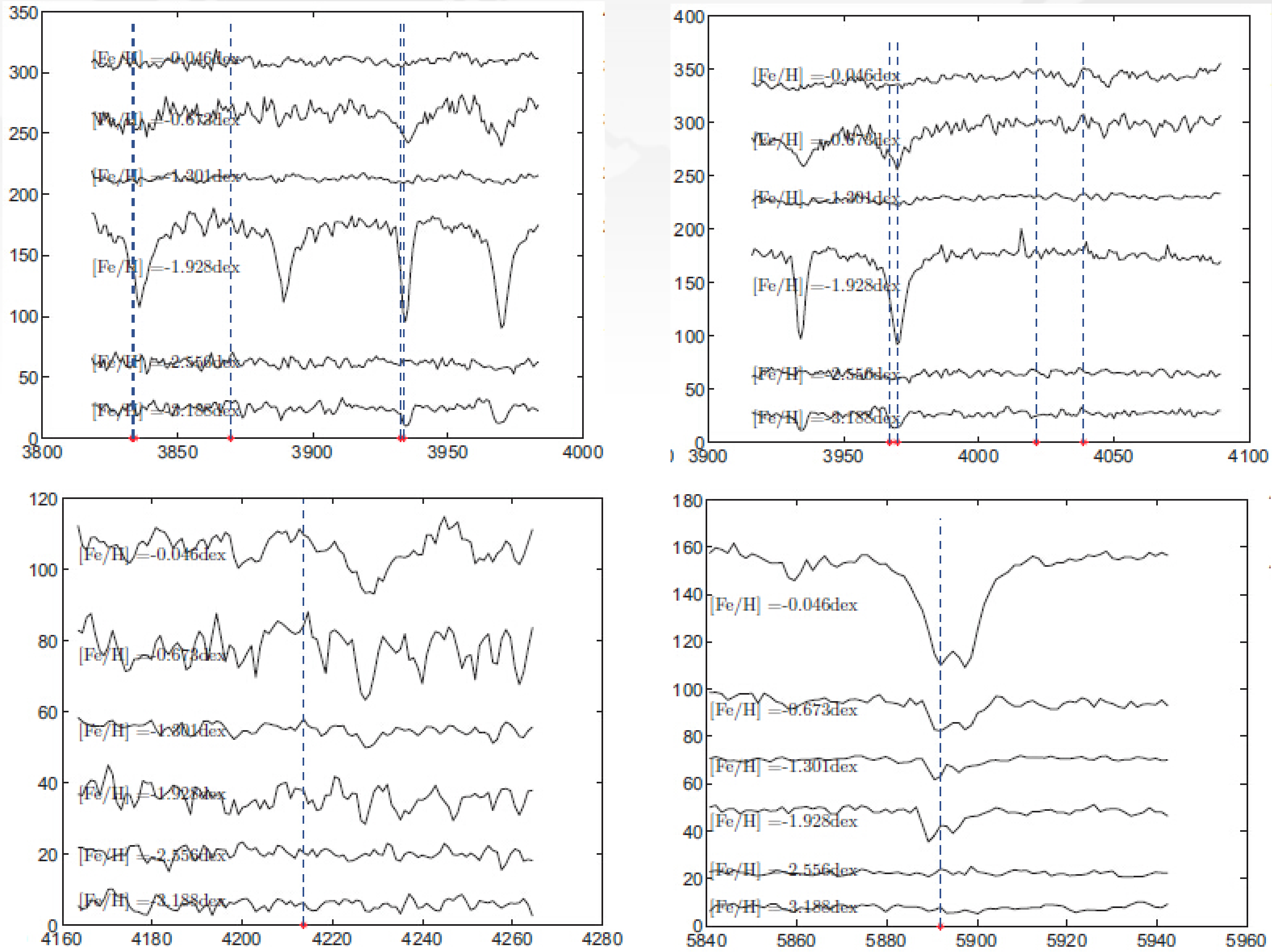
log g =4.002dex
log g =3.338dex
log g =2.675dex
log g =2.01dex
log g =1.352dex

Detection



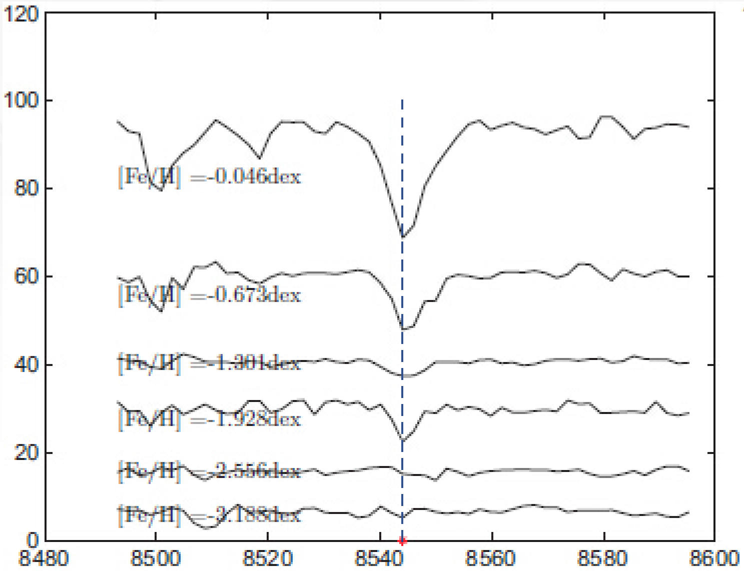
700
600
500
400
300
200
100
0
3000
4000
5000
6000
7000
8000
9000
10000
[Fe/H] =-0.046dex
[Fe/H] =-0.673dex
[Fe/H] =-1.301dex
[Fe/H] =-1.928dex
[Fe/H] =-2.556dex
[Fe/H] =-3.188dex

Detection

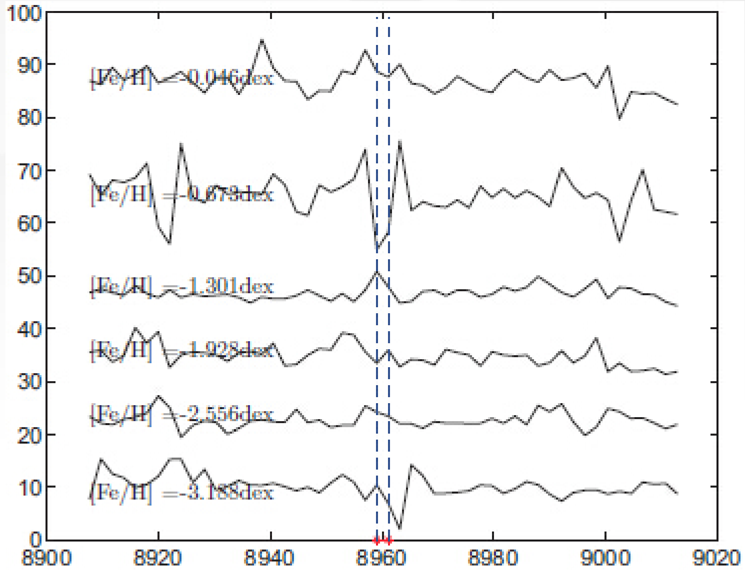


[Fe/H] =-0.046dex
[Fe/H] =-0.673dex
[Fe/H] =-1.301dex
[Fe/H] =-1.928dex
[Fe/H] =-2.556dex
[Fe/H] =-3.188dex

Detection

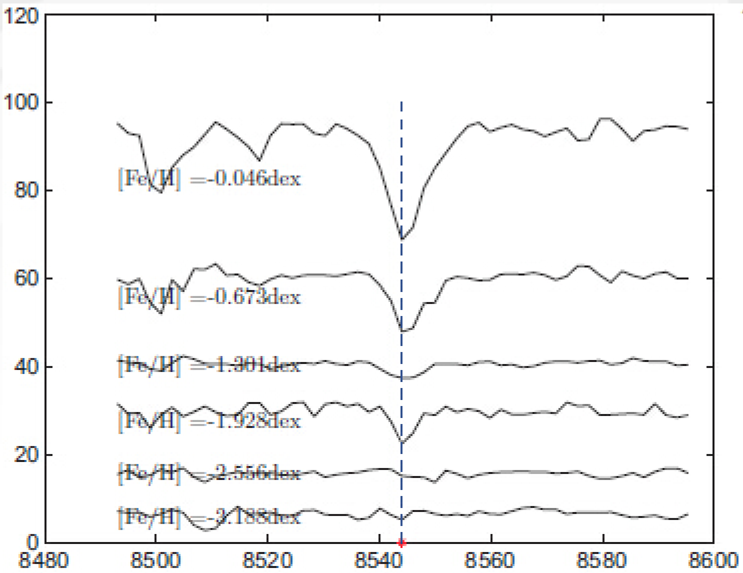


[Fe/H] =-0.046dex
[Fe/H] =-0.673dex
[Fe/H] =-1.301dex
[Fe/H] =-1.928dex
[Fe/H] =-2.556dex
[Fe/H] =-3.188dex



[Fe/H] =-0.046dex
[Fe/H] =-0.673dex
[Fe/H] =-1.301dex
[Fe/H] =-1.928dex
[Fe/H] =-2.556dex
[Fe/H] =-3.188dex

Description and Estimation



Point Description (PD)

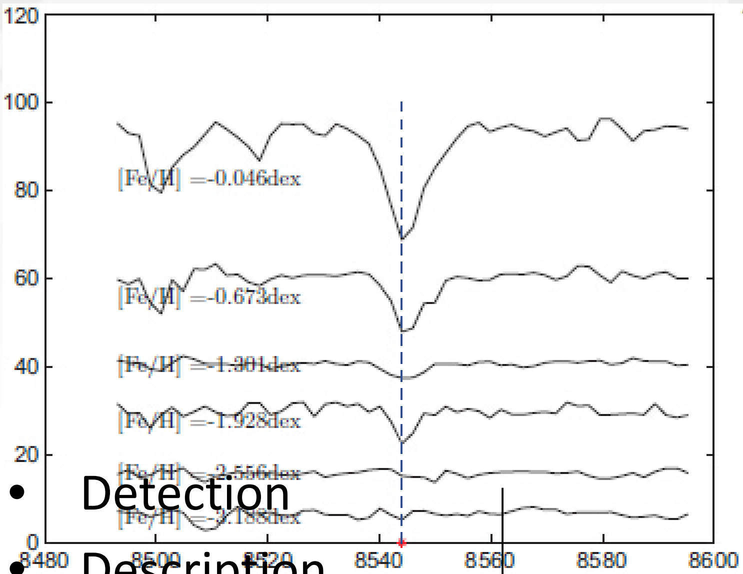
$$\beta_{i_1} \neq 0, \beta_{i_2} \neq 0, \cdots, \beta_{i_k} \neq 0$$

$$\tilde{x} = (\tilde{x}_1, \cdots, \tilde{x}_p)$$

$$x = (\tilde{x}_{i_1}, \cdots, \tilde{x}_{i_k})$$

Local Integration (LI)

Description and Estimation



- Detection
- Description
- Estimation ← f x

$$y = f(x)$$

Point Description (PD)

$$\beta_{i_1} \neq 0, \beta_{i_2} \neq 0, \cdots, \beta_{i_k} \neq 0$$

$$\tilde{x} = (\tilde{x}_1, \cdots, \tilde{x}_p)$$

$$x = (\tilde{x}_{i_1}, \cdots, \tilde{x}_{i_k})$$

Local Integration (LI)

Experimental Results

- On Real Spectra

Table 5: Consistency/Accuracy on test set with features described by the observed fluxes on the detected typical positions.

evaluation method	log $T_{\mathtt{eff}}$	log g	[Fe/H]
MAE	0.009092	0.198928	0.206814

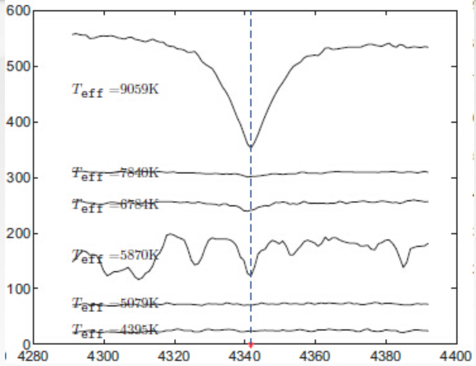


Table 6: Accuracy/Consistency on test set with features described by local integral near the detected typical positions. Integral radii are 6 for log $T_{\mathtt{eff}}$, 2 for log g and 8 for [Fe/H] respectively.

evaluation method	log Teff	log g	[Fe/H]
MAE	0.007458	0.189557	0.182060

Table 2. Mean absolute errors on the evaluation set of 19 000 spectra in the RR model (plotted in Fig. 5).

Set	PCs	+	$E_{\log\ T_{\rm eff}}$	$E_{\log\ g}$	$E_{\rm [Fe/H]}$
19 000	25 + 25		0.0126	0.3644	0.1949

Re Fiorentin, P., et al. 2007, A&A, 467, 1373

Experimental Results

- On Synthetic Spectra

Table 15
Performance on Synthetic Spectra Based on SVR and Features in Tables 12, 13, and 14

Evaluation Method	$T_{\rm eff}$	$\log g$	[Fe/H]
MAE	0.000801	0.017881	0.013142

Note. Features are described by the LI method with integration radii k = 6, 2, and 8, respectively, for $T_{\rm eff}$, $\log g$, and [Fe/H]. MAE are the mean absolute errors on the synthetic test set.

Table 1. Mean absolute errors on the evaluation set of 908 spectra in the SS model for different numbers of PCs retained in the reconstruction. (As PCA is done separately on the blue and red regions, the total number of inputs is twice the number of PCs.).

PCs	$E_{\log\ T_{\rm eff}}$	$E_{\log\ g}$	$E_{\rm [Fe/H]}$
5	0.0087	0.1264	0.1558
25	0.0036	0.0245	0.0327
100	0.0030	0.0251	0.0269
908	0.0133	0.2087	0.2308

Compactness

- On Real Spectra

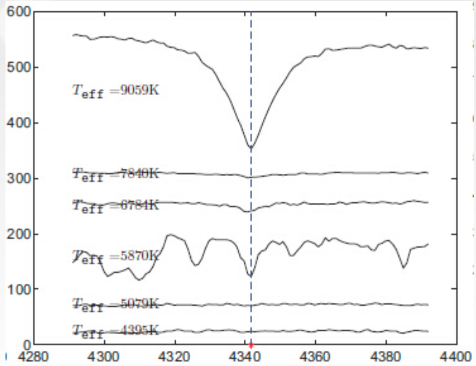


Table 5: Consistency/Accuracy on test set with features described by the observed fluxes on the detected typical positions.

evaluation method	log $T_{\mathtt{eff}}$	log g	[Fe/H]
MAE	0.009092	0.198928	0.206814

Table 6: Accuracy/Consistency on test set with features described by local integral near the detected typical positions. Integral radii are 6 for log $T_{\mathtt{eff}}$, 2 for log g and 8 for [Fe/H] respectively.

evaluation method	log Teff	log g	[Fe/H]
MAE	0.007458	0.189557	0.182060

99.74%

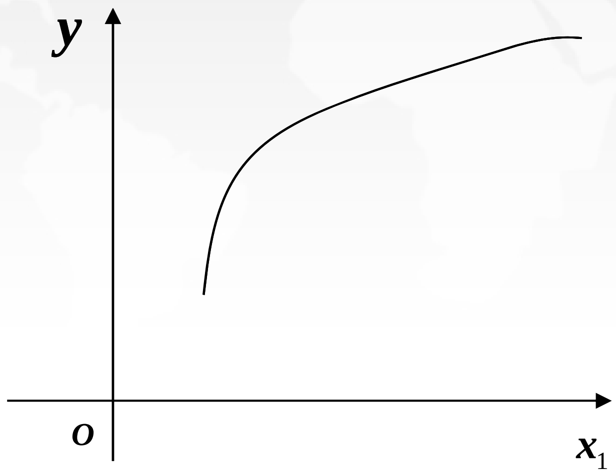
Table 2. Mean absolute errors on the evaluation set of 19 000 spectra in the RR model (plotted in Fig. 5).

Set	PCs	+	$E_{\log\ T_{\rm eff}}$	$E_{\log\ g}$	$E_{\rm [Fe/H]}$
19 000	25 + 25		0.0126	0.3644	0.1949

Re Fiorentin, P., et al. 2007, A&A, 467, 1373

Linearity v.s. nonlinearity

$$(\hat{\alpha}, \hat{\beta}) = arg \min_{(\alpha,\beta)} \{\Sigma_{i=1}^{N}(y_i - \alpha - \Sigma_{j=1}^{p}\beta_j \tilde{x}_j^i)^2\}$$



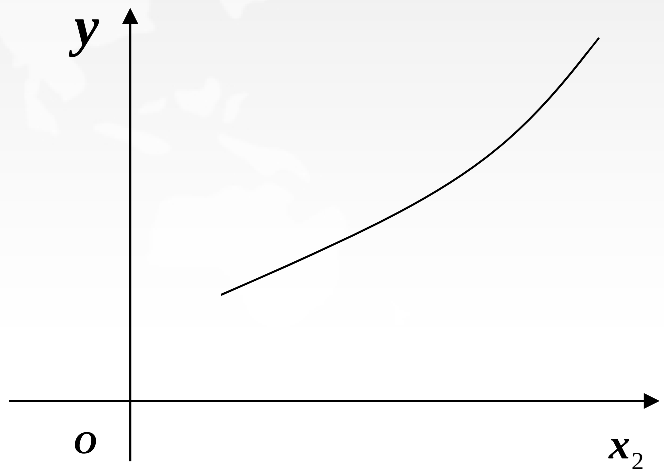


Table 10: Performance (MAE) of two linear method
Experimental configurations are same as the expe
iments in Table 6. OLS (Ordinary Least Squares
linear least squares regression, SVR(linear): Suppo
Vector machine Regresion with a linear kernel.

evaluation method	T_{eff}	$\log g$	[Fe/H]
OLS	0.036510	0.301661	0.360890
SVR(linear)	0.034152	0.253363	0.323512

Other typical non-linear estimators

Table 11: Performance (MAE) of four nonlinear methods. Experimental configurations are same as the experiments in Table 6.

evaluation method	T_{eff}	log g	[Fe/H]
FNN	0.008980	0.186014	0.179565
GAM	0.008139	0.245167	0.245111
MARS	0.011335	0.243147	0.242703
RF	0.009478	0.228717	0.204248

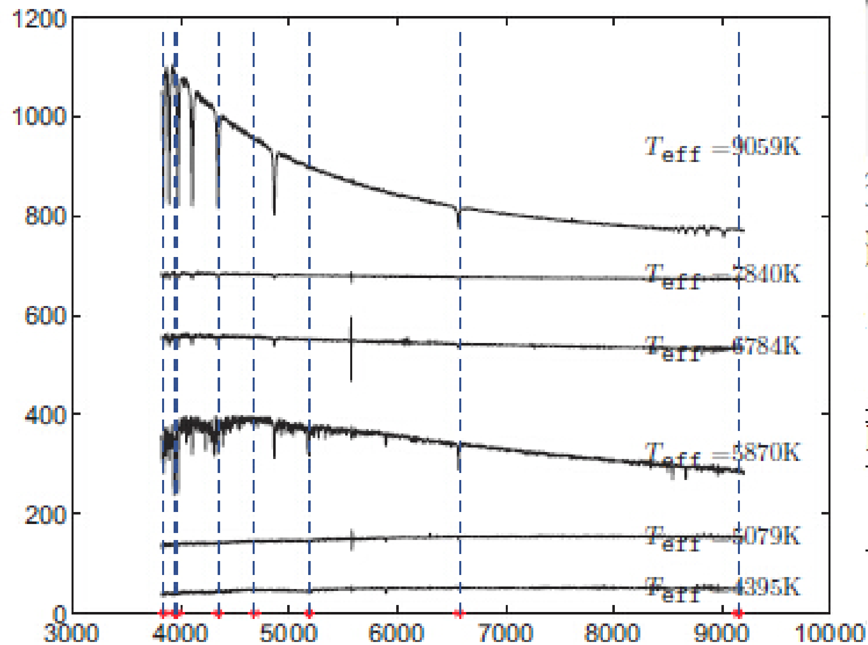
Feedforward neural network
Generalized Additive Models
Multivariate Adaptive Regression Splines
Random Forest

Table 6: Accuracy/Consistency on test set with features described by local integral near the detected typical positions. Integral radii are 6 for log $T_{\mathtt{eff}}$, 2 for log g and 8 for [Fe/H] respectively.

evaluation method	log Teff	log g	[Fe/H]
MAE	0.007458	0.189557	0.182060
SD	0.011189	0.270496	0.248504

Linearly Supporting Feature Extraction

Linearly Supporting Feature Extraction



;: Accuracy/Consistency on test set with fea-
lescribed by local integral near the detected
positions. Integral radii are 6 for log T_{eff}, 2
g and 8 for [Fe/H] respectively.

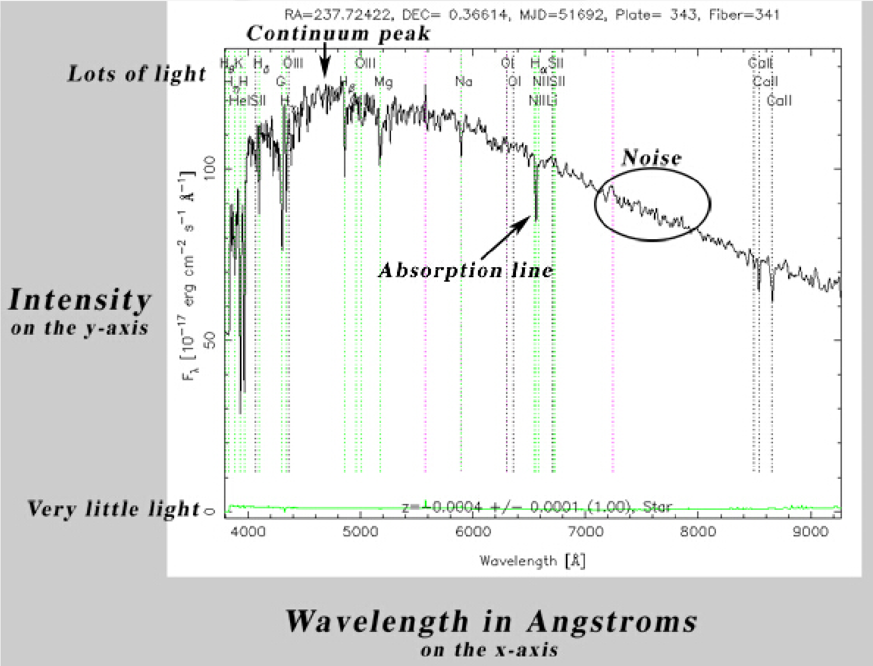
luation method	log Teff	log g	[Fe/H]
MAE	0.007458	0.189557	0.182060
SD	0.011189	0.270496	0.248504

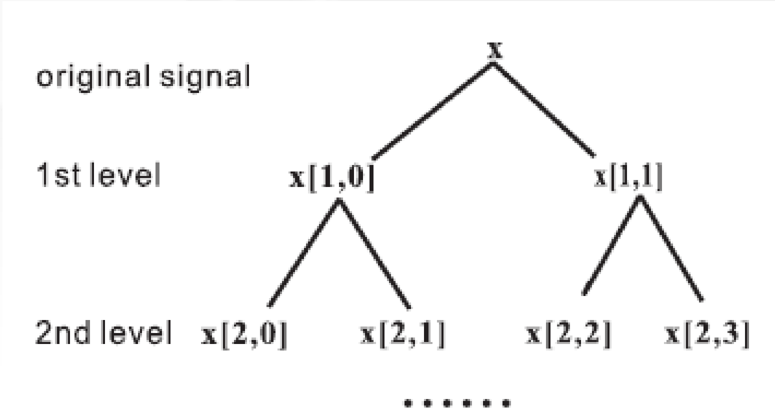
Methods

Evaluation Method	T_{eff}	$\log g$	[Fe/H]
OLS	0.036510	0.301661	0.360890
SVR(linear)	0.034152	0.253363	0.323512

$$\boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x})$$

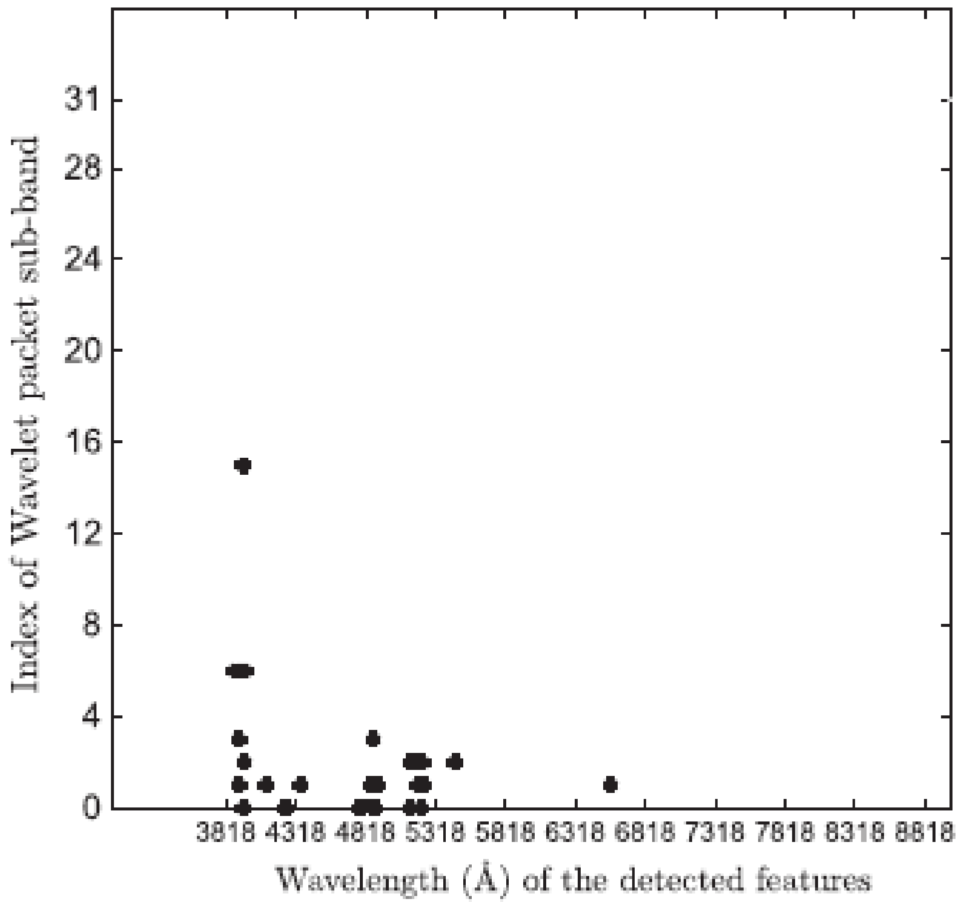
$$\hat{y} = f(\boldsymbol{x}; \boldsymbol{w}) = \sum_{j=1}^{p} w_j x_j,$$



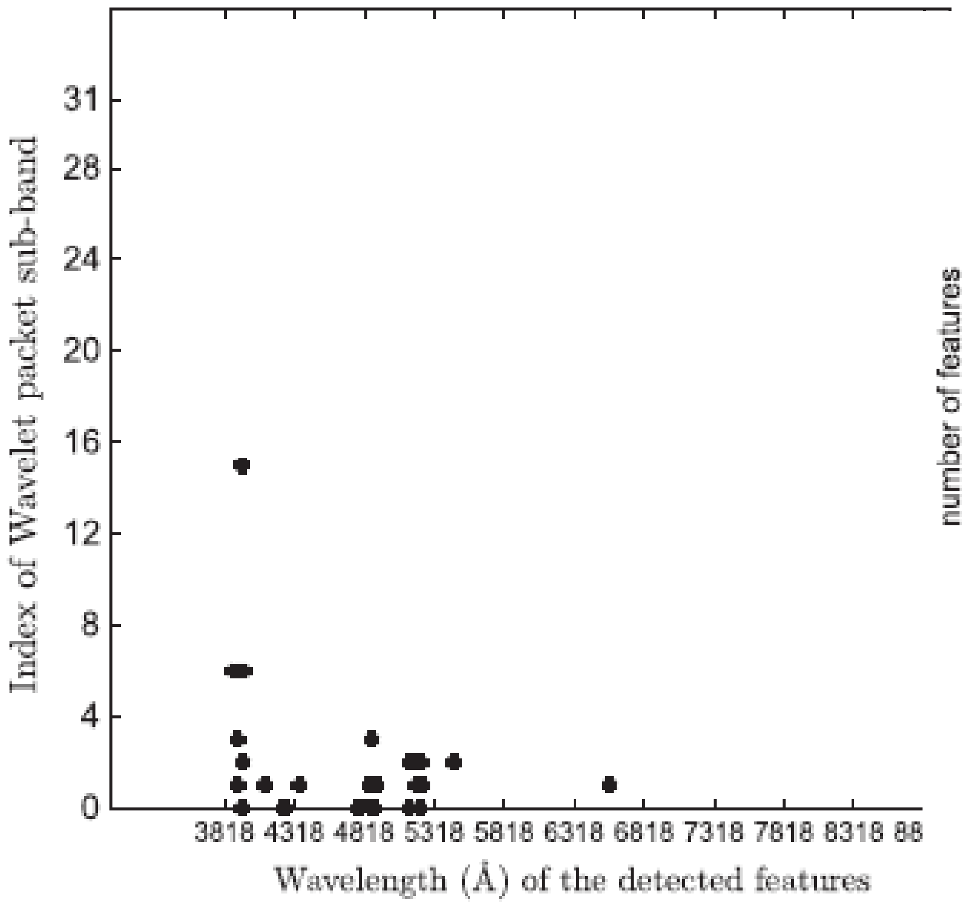


Dissolution of nonlinearity

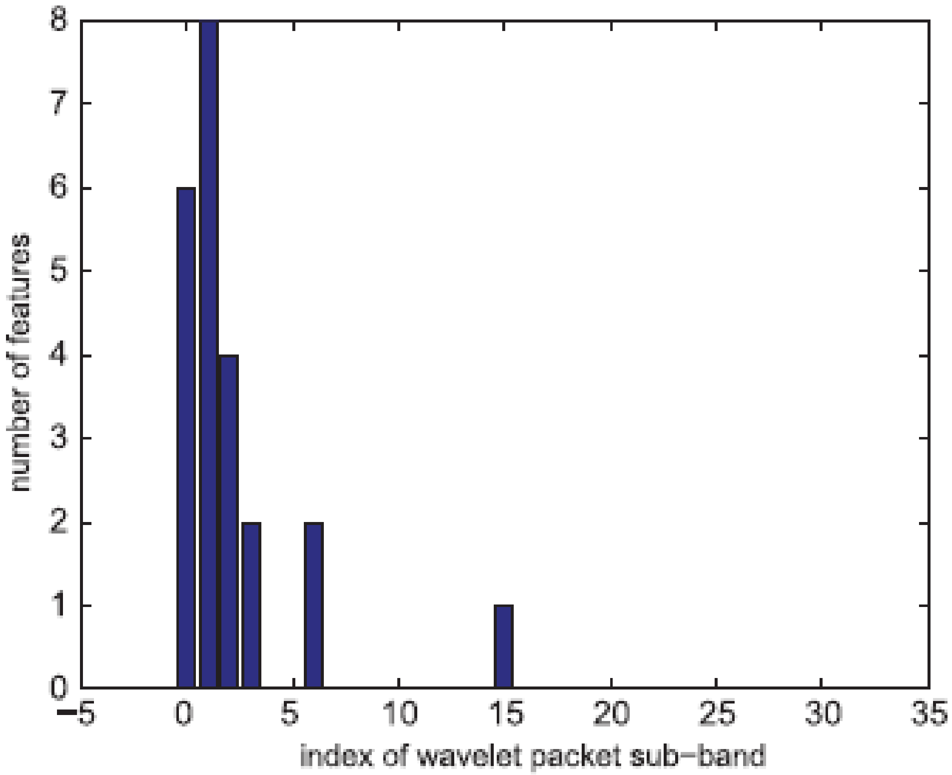
Dependeny of effectiveness on wavelength and frequency



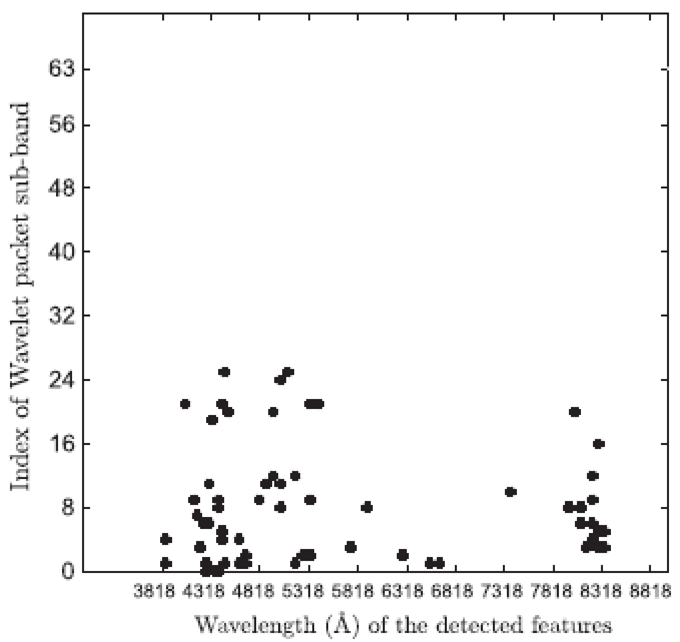
(d) Features for T_{eff}



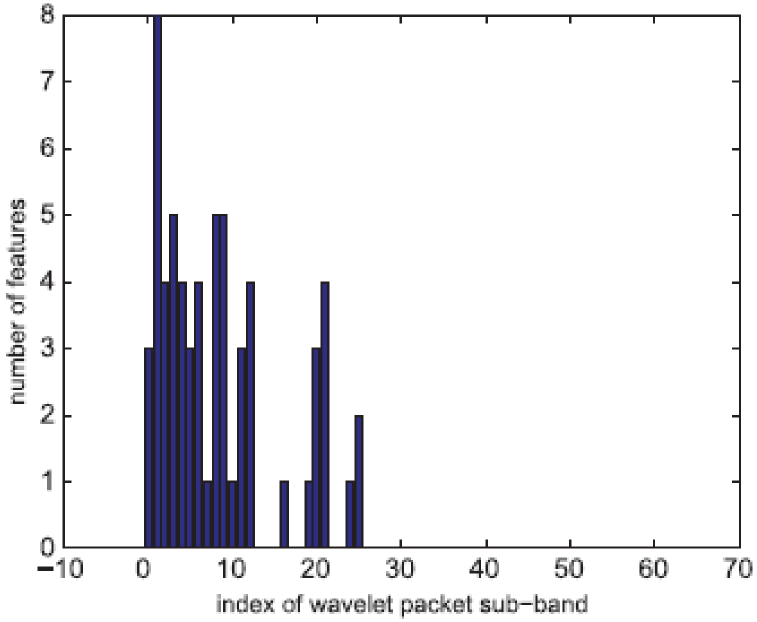
(d) Features for T_{eff}



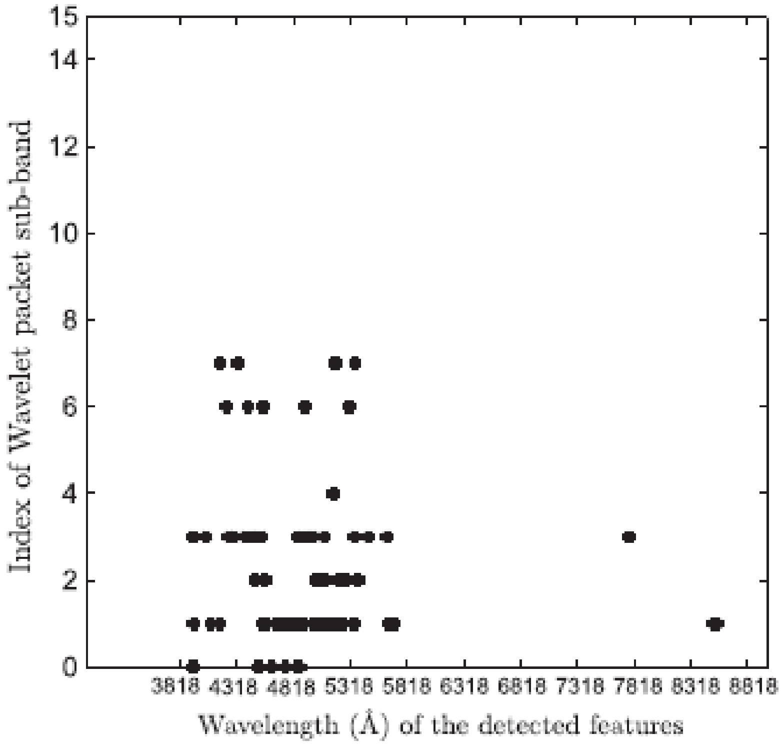
(a) Features for T_{eff}



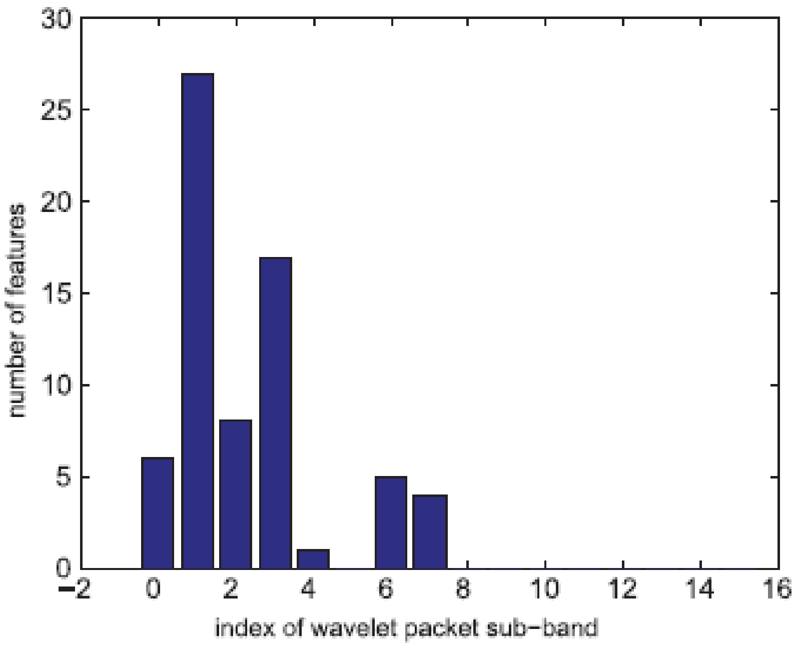
(e) Features for log g



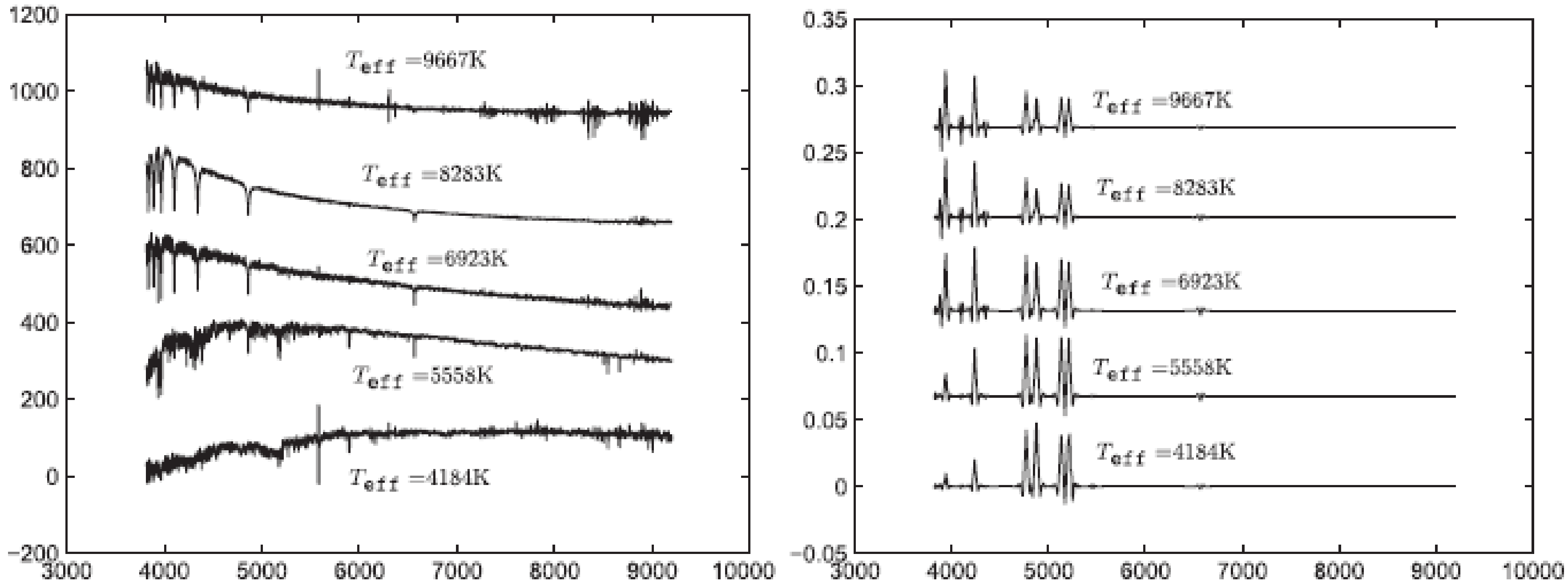
(b) Features for log g



(f) Features for [Fe/H]

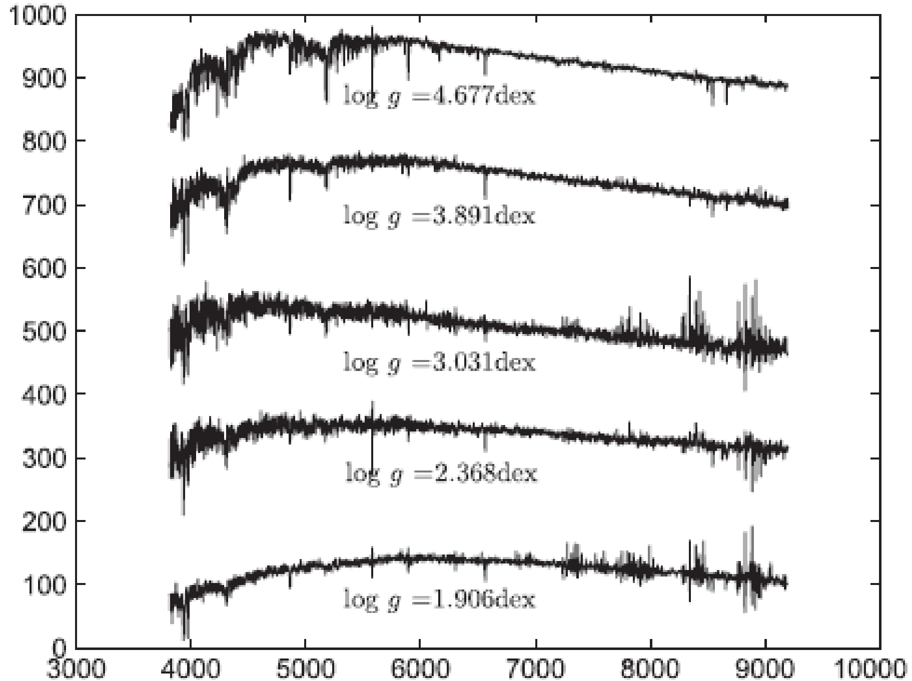


(c) Features for [Fe/H]

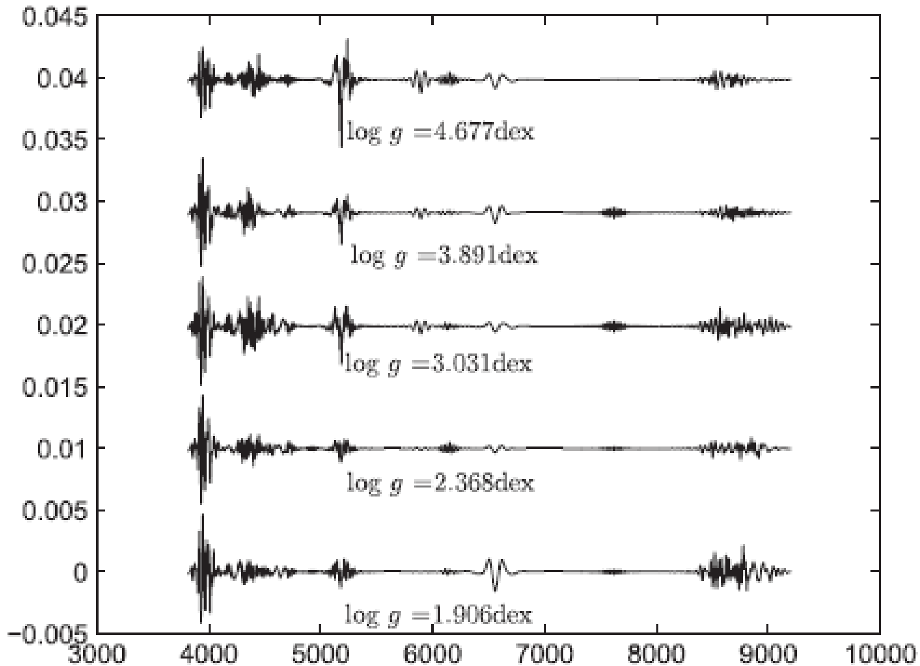


(a) Five spectra with different T_{eff}.

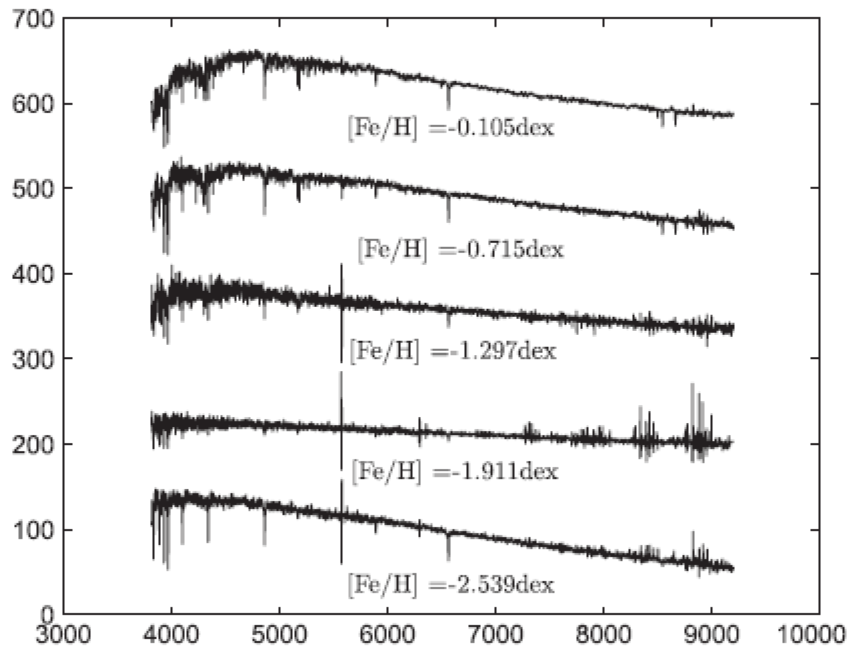
(b) Features of the spectra in Fig. 5(a).



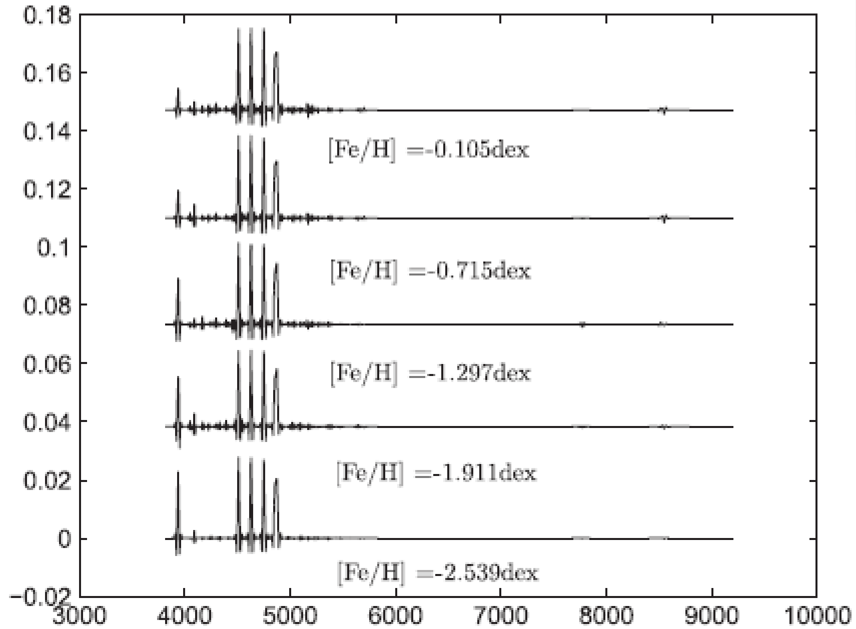
(c) Five spectra with different log g.



(d) Features of the spectra in Fig. 5(c).



(e) Five spectra with different [Fe/H].



(f) Features of the spectra in Fig. 5(e).

(a) The Detected Features for T_{eff} Based on Basis Function rbio with the Optimal Decomposition Level 5

Label	TW λ(Å)	IF	label	TW λ(Å)	IF
T_1	[3825.6,3936.4,4050.4]	0	T_2	[4118.1,4237.4,4360.1]	0
T_4	[4737.0,4874.2,5015.3]	0	T_5	[4987.7,5132.2,5280.8]	0
T_7	[3818.6,3903.9,3991.2]	1	T_8	[3998.5,4099.2,4202.4]	1
T_{10}	[4737.0,4856.2,4978.5]	1	T_{11}	[4772.0,4892.2,5015.3]	1
T_{13}	[5099.2,5227.6,5359.2]	1	T_{14}	[6407.7,6569.0,6734.4]	1
T_{16}	[5006.1,5141.6,5280.8]	2	T_{17}	[5080.4,5218.0,5359.2]	2
T_{19}	[3818.6,3903.9,3991.2]	3	T_{20}	[4754.4,4865.2,4978.5]	3
T_{22}	[3846.8,3947.3,4050.4]	6	T_{23}	[3850.3,3934.6,4020.7]	15

label	TW λ(Å)	IF
T_3	[4633.4,4767.6,4905.7]	0
T_6	[5061.7,5208.3,5359.2]	0
T_9	[4241.3,4348.1,4457.6]	1
T_{12}	[5061.7,5189.2,5319.9]	1
T_{15}	[3839.7,3943.7,4050.4]	2
T_{18}	[5310.1,5453.8,5601.4]	2
T_{21}	[3818.6,3875.3,3932.8]	6

For log g, [Fe/H] …

$$\hat{y} = f(x; w) = \sum_{j=1}^{p} w_j x_j,$$

Label	w_i	label	w_i	label	w_i	label	w_i
T_1	0.2379	T_2	0.4505	T_3	0.4819	T_4	−1.8345
T_8	0.4558	T_9	0.7899	T_{10}	1.0640	T_{11}	−0.9495
T_{15}	0.2825	T_{16}	0.5646	T_{17}	0.5911	T_{18}	−0.4539
T_{22}	0.2366	T_{23}	0.4507	…	…	…	…

label	w_i	label	w_i	label	w_i
T_5	1.0740	T_6	0.7236	T_7	0.3264
T_{12}	0.8358	T_{13}	0.8650	T_{14}	1.5004
T_{19}	0.3262	T_{20}	−0.9225	T_{21}	−0.2100

Label	w_i	label	w_i	label	w_i	label	w_i
L_1	−14.9500	L_2	11.9097	L_3	−15.9727	L_4	−29.8595
L_8	124.5594	L_9	−23.4959	L_{10}	4.2752	L_{11}	9.3749
L_{15}	20.5712	L_{16}	26.5010	L_{17}	32.2221	L_{18}	11.0103
L_{22}	−6.9629	L_{23}	−10.0933	L_{24}	−11.4390	L_{25}	−26.8052
L_{29}	16.4676	L_{30}	−9.1541	L_{31}	−47.8560	L_{32}	−7.0705
L_{36}	−43.4952	L_{37}	−11.2162	L_{38}	−12.9840	L_{39}	8.6831
L_{43}	−19.2899	L_{44}	−9.5705	L_{45}	7.2348	L_{46}	−12.2654
L_{50}	−16.5275	L_{51}	−8.4650	L_{52}	−14.5751	L_{53}	−19.9664
L_{57}	−3.6372	L_{58}	−28.5637	L_{59}	−8.3460	L_{60}	38.7862

label	w_i	label	w_i	label	w_i
L_5	20.9904	L_6	−12.1664	L_7	48.9051
L_{12}	−11.0132	L_{13}	−26.4751	L_{14}	−35.4924
L_{19}	−22.3564	L_{20}	11.6441	L_{21}	−14.6373
L_{26}	−25.4320	L_{27}	9.7027	L_{28}	−21.6698
L_{33}	−18.4181	L_{34}	−19.0691	L_{35}	−19.7537
L_{40}	13.4435	L_{41}	−13.0983	L_{42}	−5.3274
L_{47}	16.4605	L_{48}	−14.0761	L_{49}	34.1498
L_{54}	−28.4143	L_{55}	−17.2212	L_{56}	18.7294
L_{61}	−15.1417	L_{62}	−23.3872	…	…

Label	w_i	label	w_i	label	w_i	label	w_i
F_1	−15.9393	F_2	14.4512	F_3	10.0133	F_4	14.5021
F_8	−3.5740	F_9	−4.2593	F_{10}	15.7326	F_{11}	15.0760
F_{15}	−16.4978	F_{16}	−19.8914	F_{17}	−6.1936	F_{18}	−13.9105
F_{22}	14.8693	F_{23}	−6.8806	F_{24}	14.3556	F_{25}	−19.2167
F_{29}	−7.0540	F_{30}	13.4940	F_{31}	−13.5002	F_{32}	−8.6885
F_{36}	−12.3663	F_{37}	5.2848	F_{38}	6.5740	F_{39}	−21.8866
F_{43}	4.5351	F_{44}	−5.7992	F_{45}	7.4311	F_{46}	1.2759
F_{50}	13.3047	F_{51}	−11.3030	F_{52}	8.5673	F_{53}	7.0954
F_{57}	−16.7121	F_{58}	−10.0843	F_{59}	−17.2424	F_{60}	−2.3355
F_{64}	15.6214	F_{65}	4.5253	F_{66}	−6.8905	F_{67}	17.1603

label	w_i	label	w_i	label	w_i
F_5	−36.4275	F_6	−22.7435	F_7	−20.0956
F_{12}	9.5922	F_{13}	15.1001	F_{14}	−10.5834
F_{19}	−9.9439	F_{20}	−16.1353	F_{21}	10.6011
F_{26}	−43.5510	F_{27}	3.8668	F_{28}	−19.1643
F_{33}	9.5827	F_{34}	17.6601	F_{35}	5.7138
F_{40}	−23.7674	F_{41}	−16.2051	F_{42}	−10.6150
F_{47}	−8.2072	F_{48}	5.1707	F_{49}	1.6828
F_{54}	6.1268	F_{55}	15.5904	F_{56}	−5.3280
F_{61}	−4.8779	F_{62}	−8.4915	F_{63}	10.0789
F_{68}	−8.5783	⋯	⋯	⋯	⋯

		(a) Comparison with SDSS Data Set			
Estimation Method	Evaluation Method	log T_{eff} (dex)	T_{eff} (K)	log g (dex)	[Fe/H](dex)
Linear:OLS	MAE	0.0062	82.94	0.2345	0.1564
	ME	0.0002	2.769	−0.0219	−0.0003
	SD	0.0096	135.9	0.3297	0.2196
Nonlinear:ANN [1]	MAE	0.0126	⋯	0.3644	0.1949
Nonlinear:MAχ [2]	ME	⋯	130	0.5	0.24
Nonlinear:SVR$_G$[3]	MAE	0.007458	101.610	0.189557	0.182060
Linear:OLS [4]	SD	⋯	196.473	0.596	0.466
Linear:SVR$_l$ [3]	MAE	0.034152	⋯	0.253363	0.323512

	(b) Comparison with Synthetic Data Set Derived from Kurucz's NEWODF Models (Castelli & Kurucz 2003)				
Estimation Method	Evaluation Method	log T_{eff} (dex)	T_{eff} (K)	log g (dex)	[Fe/H](dex)
Linear:OLS	MAE	0.0022	31.70	0.0337	0.0268
Nonlinear:ANN [1]	MAE	0.0030	⋯	0.0245	0.0269

Note. OLS (ordinary least squares): linear least squares regression, SVR$_l$: support vector machine regression with a linear kernel, SVR$_G$: suppo regression with a Gaussian kernel, ANN: Artificial Neural Networks, MAχ: MAssive Compression of χ^2. (1):Re Fiorentin et al. (2007), (2):Jofre e et al. (2014), (4):Tan et al. (2013)

Performance of the proposed scheme on 23963 test spectra from LAMOST (10000 LAMOST spectra for training, Section 2.2).

method	log Teff(Teff)		
parameter	MAE	ME	SD
RBFNN	0.0070(91.14)	6.75×10^{-5}(2.37)	0.0099(131.36)
SVR_G	0.0074(95.37)	1.27×10^{-4}(4.30)	0.0106(141.62)
KNNR	0.0085(111.80)	4.47×10^{-4}(10.08)	0.0126(173.66)
LSR	0.0082(106.51)	6.67×10^{-4}(11.46)	0.0117(161.83)
SVR_l	0.0081(105.73)	1.89×10^{-4}(5.44)	0.0116(159.83)

method	log g			[Fe/H]		
parameter	MAE	ME	SD	MAE	ME	SD
RBFNN	0.1664	0.0109	0.2753	0.1197	-0.0038	0.1767
SVR_G	0.1528	-0.0008	0.2102	0.1146	-0.0112	0.1528
KNNR	0.1934	0.0167	0.2730	0.1625	-0.0154	0.2151
LSR	0.2218	0.0173	0.3404	0.1312	-0.0143	0.1807
SVR_l	0.2070	-0.0124	0.3145	0.1311	-0.0222	0.1806

Thanks